

EXHIBIT 73

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/228757975>

Improving the quality of retrospective reports: Calendar interviewing methodologies

Article · January 2004

CITATIONS

5

READS

697

1 author:



Robert Belli

University of Nebraska–Lincoln

102 PUBLICATIONS 4,411 CITATIONS

SEE PROFILE



Working Paper series of the Program in Survey Research and Methodology
University of Nebraska, Lincoln
Gallup Research Center
<http://sram.unl.edu>

Improving the Quality of Retrospective Reports: Calendar Interviewing Methodologies

Robert F. Belli

Program in Survey Research and Methodology, University of Nebraska, Lincoln

Working paper no 12
September 2005

This version is forthcoming on the proceedings of the 2004 RC33 conference
<http://www.siswo.uva.nl/rc33/>

The full citation is: Beelli, R. F. (forthcoming). Improving the Quality of Retrospective Reports: Calendar Interviewing Methodologies. In C. van Dijkum, J. Blasius, H. Kleijer & B. van Heiten (Eds.), *Proceedings of the RC 33 Sixth International Conference on Social Science Methodology. Recent Developments and Applications in Social Research Methodology (CD-ROM)*. Wiesbaden, Germany: VS Verlag.

Corresponding author: Robert F. Belli
email: bbelli2@unlnotes.unl.edu

Improving the Quality of Retrospective Reports: Calendar Interviewing Methodologies

Robert F. Belli
UNL—Gallup Research Center
Department of Psychology
University of Nebraska—Lincoln
bbelli2@unl.edu

Selected Session: Cognitive Aspects of Survey Methodology and Their Applications (I): Aided Recall Techniques in Survey Interviews

Acknowledgements: This research has been supported in part by the National Institute of Health (National Institute on Aging and National Institute of Child Health and Human Development) grant 1R01AG17977, and by the National Science Foundation grant SES-0001994.

INTRODUCTION

Social, economic, and health scientists have been increasingly using cross-sectional and panel survey methods to identify factors that contribute to quality of life. Such methods often rely on retrospective behavioral reports provided at the time of interviewing to identify influential variables in respondent histories. Despite its widespread use, research on human memory suggests that the quality of data obtained through retrospective reporting is seriously compromised by the limitations of respondent memory. Specifically, past research indicates that the retrospective reporting of events using traditional question-list (Q-list) survey methods results in greater response error, especially for events that occur more remotely in time. This problem seriously compromises the validity of any inferences generated from data relying on respondent recall. Fortunately, advances in cognitive science suggest that emerging Event History Calendar (EHC) methodologies improve the quality of retrospective reports.

Overview of Retrospective Reporting Errors with Question-List (Q-list) Methods

Use of traditional Q-list approaches results in retrospective reporting errors that become considerably more pronounced with events that have occurred in the more distant past. This problem is evident in reports of key health and economic variables that occurred just one to two years previous to the interview. For example, Cannell et al. (1965) found in a record check study that respondents failed to report only 3% of hospitalizations when asked within 1 to 10 weeks of the event, but 42% of hospitalizations were not reported when asked 1 year after the event. Belli and Lepkowski (1996) also found better data quality in the reporting of doctor's office visits with a 6-month recall in comparison to a 12-month recall. With regard to economic variables, 2-year recall periods have resulted in poorer quality reports than 1-year recall periods by showing 1) greater bias and error variance in the reporting of annual and hourly earnings, and hours spent in employment (Duncan & Hill, 1985), 2) a reduction in the accurate reporting of those months in which

unemployment occurred (Mathiowetz, 1986; Mathiowetz & Duncan, 1988), and 3) a reduction in the accurate reporting of occupational change (Duncan & Mathiowetz, 1985). Similarly, in a longitudinal panel study, Dugoni, Lee, and Tourangeau (1996) compared the reporting for a reference period that occurred between 1 to 2 years ago against the reports for the same period of time provided from these same respondents one-year earlier. They found that the two-year recall led to fewer reports of number of jobs and whether unemployment compensation was received in comparison to the one-year recall reports. This memory performance issue poses serious threats to any inferences that are generated from retrospective survey data and is also problematic for panel intake surveys, in which respondents are asked to retrospectively report on events that have occurred over their entire life course.

The Potential of Event History Calendars

Problems with the quality of retrospective reports obtained with Q-list interviews may be ameliorated with new methodologies that a) optimize usage of autobiographical memory cues and b) allow for a more narrative or conversational style of respondent expression. An Event History Calendar (EHC) questionnaire methodology has considerable potential toward improving the quality of retrospective reports in comparison to Q-list methods. EHCs may promote a greater completeness and accuracy in the reporting of spells and transitions in respondents' pasts, and they may provide a framework for more conversational and motivating interactions between interviewers and respondents.

EHCs and Response Accuracy. The use of EHCs or charts to collect retrospective information has already had a wide impact on the fields of population studies and sociology (Axinn et al., 1997; Caspi et al., 1996; Freedman et al., 1988; Furstenberg et al., 1987; Kessler & Wethington, 1991; Lyketsos et al., 1994; van der Vaart, in press) and their use has demonstrated high quality retrospective reporting. For example, Freedman et al. (1988) compared the retrospective reports of school attendance and weekly work hours in the life history calendar against the concurrent reports gathered with Q-list interviewing that had taken place 5 years earlier. In examining the same month of the prior concurrent reports, 87% of respondents gave identical answers in both of the interviews regarding full-time, part-time, or no school attendance. In another application of a life history calendar, Caspi et al. (1996) found at least 90% agreement between retrospective reports of activities for a given month on the life history calendar and concurrent reports obtained 3-years earlier with living arrangements, cohabitation, schooling, employment, and job training.

Theoretically, EHCs acquire their ability to optimize the quality of retrospective reports because they tap into available idiosyncratic structures in autobiographical memory (Belli, 1998). In contrast to Q-lists, which emphasize structuring the interview through standardized scripted questions, EHCs are visually structured instruments that emphasize the interrelationships of spells of activity in one's autobiographical past by presenting timelines to record what events happened, and when. In order to tap into idiosyncratic autobiographical memory structures, EHCs encourage a flexible style of interviewing in which events from respondents' pasts are used as cues to facilitate the recall of other

related events. These cueing mechanisms optimize the ability of respondents to reconstruct an accurate and complete memory of the occurrence and timing of past events, even after retention intervals of many years. Moreover, these cueing mechanisms are largely ignored in Q-list instruments; such a deficit may explain the degree of retrospective reporting error found with Q-list questionnaires (Belli et al., in press). As well, the presence of cueing strategies in EHCs should facilitate enhanced recall for life events that have occurred in the more distant past compared to more recent life events. Although this hypothesis is yet to be tested, verbal behavior analyses of the life course data in the proposed research is expected to support this hypothesis.

EHCs and Conversational Patterns. The benefits of EHC interviewing go beyond those simply associated with memory cueing. EHCs allow for more effective communication patterns, in comparison to standardized interviewing techniques. EHCs promote a narrative style of remembering which is compatible with the manner in which autobiographical knowledge is structured (Brown & Schopflocher, 1998; Schank & Abelson, 1995). A behavioral coding analysis of the 2-year paper-and-pencil PSID Calendar Methods Study provided initial support for this claim; after identifying verbal interaction patterns, it was found that EHC interviewing produced greater conversational engagement than Q-list methods, as indicated by higher levels of clarification and feedback behaviors (Belli et al., in press). Note also that despite the EHC collecting more detailed information than the Q-list, respondents reported the same levels of burden for both methods. As well, these differences in data quality were not attributable to differences in interviewers' abilities to elicit quality responses, as revealed by analyses that tested for differences in interviewer variance between EHC and Q-list conditions. Although promising, similar analyses have yet to be carried out on *computer-assisted* EHC and Q-list *life course* instruments.

The results reported in the present paper review three studies that have directly compared the quality of retrospective reports obtained through traditional standardized Q-list and the more flexible style of EHC interviewing. Study 1 examines paper and pencil instruments administered via telephone, and examines a reference period of approximately 1 to 2 years previously. Both Studies 2 and 3 examine the quality of retrospective reporting for a life course reference period. Study 2 examines paper and pencil instruments administered face-to-face, and Study 3 examines computer-assisted-interviewing instruments administered via telephone. In all three studies, EHC interviews are found to promote higher quality retrospective reports in comparison to Q-list interviews, across different types of instruments and interviewing modes.

STUDY 1

Method

A more detailed presentation of methods in Study 1 are presented by Belli, Shay, and Stafford (2001). Interviews were collected in 1998 via telephone on paper with a random subset of respondents and questions from the regular 1997 PSID core survey. Two conditions were tested, a Q-list condition ($n = 307$; 84.1% response rate), and an EHC

condition ($n = 309$; 84.4% response rate). Respondents and 20 interviewers were randomly assigned to conditions. All interviewers received general interviewer training plus 15 hours (5 hours for each of three days) of training in their respective methods. The Q-list condition used a traditional standardized survey instrument based on the core PSID questionnaire, a 25-page paper questionnaire with scripted question sequences. Interviewers were asked to optimize retrospective reporting within the constraints of standardized interviewing techniques. The EHC condition employed paper abstracts from the Q-list method as reference material, but the survey instrument, per se, was one 18"x 28" page; flexibility in using sequential and parallel cues were facilitated by the layout of timelines within domains in the instrument. With the EHC, interviewers were given scripted questions to introduce each domain and were instructed on how to use top-down, sequential, and parallel cues. Interviews were conducted in 1998 about events that occurred during 1996 and 1997.

Results

Using data from respondents collected one year earlier on events reported during the 1997 core PSID as a standard of comparison, the quality of retrospective reports on 1996 events from the 1998 administration of EHC and Q-list interviews was assessed. The results of the experiment supported the notion that EHC methods promote a better utilization of beneficial autobiographical memory processes in comparison to state-of-the-art standardized Q-list methods. Better quality retrospective reporting for residence changes, number of jobs, earned income, weeks unemployed, weeks away from work due to personal illness and illness of another, was found with the EHC interviews for retrospective reports that targeted a reference period that was one to two years ago. There were no substantive differences in interviewing time between methods. Follow-up questionnaires revealed that interviewers enjoyed the EHC interviews more than the Q-list ones, they found the EHC easier to administer, and that they believed that the respondents found the questions on the EHC to be easier to answer. Despite the EHC collecting much more detailed information than the Q-list, respondents reported the same levels of burden for both methods.

The benefits of EHC interviewing in comparison to Q-list methods is illustrated by examining the marginal results of reports for residential moves and number of jobs treated categorically. Consider the proportion of respondents who reported moving at least once during 1996 or early 1997, and the proportions who reported having 0, 1, or 2 or more jobs during 1996, when responding in the 1997 core PSID standard of comparison interviews and in the 1998 experimental interviews. Figure 1 depicts the results for reported moves, using two different standard of comparison assessments. With the yes/no standard of comparison, responses of "yes" or "no" to a question on the 1997 core PSID interview on whether respondents reported having moved since the spring of 1996 was used. With the concurrent reports of address as a standard of comparison, the addresses that were concurrently reported by respondents during the 1996 and 1997 regular annual interviews were used; if the addresses were the same, the respondent was determined to not have moved during the overlap period, and if they were not the same, the respondent was determined to have moved. Figure 2 depicts the proportions of

respondents on reports of number of jobs. Both figures illustrate the overall tendency in the experimental conditions to underreport relative to the standard of comparison reports. The figures also show that this tendency to underreport is less pronounced in the EHC condition relative to the Q-list condition.

With reported moves (see Figure 1), and the yes/no responses as standard of comparison, the Q-list condition underreported moves by a level of 6.2% (from 18.6% in the standard of comparison to 12.4% in the Q-list) whereas the EHC condition only led to a 1.6% level of underreporting (from 20.5% to 18.9%). Using the concurrent address reports as a standard of comparison, the data pattern with moves is replicated, with the Q-list condition leading to a 8.1% reduction in reported moves (from 20.7% to 12.6%), and the EHC condition resulting in a 4.3% reduction (from 23.0% to 18.7%).

In reporting the number of jobs held during 1996, the Q-list condition underreported the proportion of respondents who reported having held two or more jobs in 1996 by 9.4% (from 24.4% to 15.0%; see Figure 2), and overreported the proportion of respondents who held one job in 1996 by 8.2% (from 51.8% to 60.0%). Overall, then, there is a disproportionate reduction in the number of jobs reported in the Q-list condition relative to the standard of comparison. These results contrast sharply with the EHC condition, which reproduced the proportions found in the standard of comparison for reports of 0, 1, and 2 or more jobs, almost exactly. This finding is especially crucial given the focus of PSID on labor economics.

Figure 3 illustrates the substantive degree of differences in the correlations between experimental and standard of comparison reports that were found for earned income, number of months employed, number of weeks unemployed, months unemployed, weeks away from work due to personal illness, weeks away from work due to illness of another, and months in receipt of ADC. All of these correlations, with the exception of months of ADC, are significantly stronger in the EHC condition. In total, there were 23 comparisons between EHC and Q-list measures. Of these, the EHC led to higher quality retrospective reports in 12 comparisons, and the Q-list in only one of them (months of ADC). In comparison to the Q-list, the EHC was found to provide better correspondence with the standard of comparison for whether one moved, the number of jobs, income, the number of months employed, the number of weeks and months unemployed, the number of months out-of-the labor force, the number of weeks missing work due to personal illness and the illness of another, which months employed, and which months unemployed. The advantages that the EHC condition reveal in comparison to Q-list interviewing are even more remarkable when considering that the format of interviewing in the EHC condition was less similar to the standard of comparison than that engendered in the Q-list conditions, as the standard of comparison itself was derived from a Q-list interviewing methodology.

STUDY 2

Method

Study 2 is a comparative study that has been conducted by Yoshihama et al. Retrospective reports of intimate partner violence (IPV) from two groups were compared. The EHC group consisted of 40 low income single African American mothers aged 18-54 years. The Q-list group consisted of 359 respondents from the Mothers' Well-being Study (MWS) who matched the EHC group in demographics and who came from the same sampling frame. Both groups were administered questions in face-to-face format

Before being asked the IPA questions, respondents in the Q-list condition were asked about demographic characteristics, their health conditions, mental health symptoms, experiences with substance use, health care utilization, any medications being taken, and levels of social support.. The IPA questions themselves asked whether respondents ever experienced 22 behavior-specific types of IPA during their life time. Including 12 types physical violence, 2 types sexual violence, and 8 types threats and harassment. For each IPA type reported, respondents were then asked whether they had experienced this type of APA in the past 12 months. Finally, respondents were asked the age at which any type of IPA first experienced.

In the EHC condition, before being asked IPA questions, respondents were exposed to timeline domains for purposes of reporting their residential history, schools attended, work history, births of offspring, experiences with financial difficulty, and relationship history (partner's initials, level of sexual behavior, cohabitation, marriage, separation, divorce). The same 22 types of IPA as in MWS were asked, but as an EHC timeline domain.

Results

Data were analyzed using survival analysis methodology. The results demonstrate superior performance of EHC interviewing in two ways. First, as shown in Table 1, the EHC led to more numerous reports of all types of IPA in comparison to the Q-list; as experiences with IPV are likely to be underreported, the greater number of reported experiences in the EHC condition are evidence of higher quality retrospective reports. Second, the Q-list condition revealed an age-cohort effect, whereas the EHC condition did not. Specifically, in the Q-list condition, older women were reporting their first experience with IPV at significantly older ages than younger women. In the EHC condition, older and younger women reported first experiences with IPV at the same ages. Apparently, the older women in the Q-list condition failed to remember their earliest experiences with IPV, whereas older women in the EHC condition were able to remember and report their earliest experiences.

STUDY 3

Method

Interviews were collected during three months of interviewing between July and September of 2002 via telephone, with a random subset of respondents from the regular 2001 Panel Study of Income Dynamics (PSID) core survey. Eligible respondents for the study were PSID study participants since at least 1980, who were respondents for at least ½ of the years they were in the study, and were at least 45 years of age. There were two conditions, a Q-list condition (N=315, 96% cooperation rate, AAPOR definition #1) and an EHC condition (N=311, 93% cooperation rate). Respondents were randomly assigned to one of the two conditions. Twenty-eight interviewers were first matched on experience and then randomly assigned to one of the two conditions. Respondents were offered \$50.00 as a token of appreciation for their participation.

The Q-list condition used a traditional computer assisted standardized survey instrument programmed in Blaise®. The EHC condition used a computer-assisted calendar, copyright (c) 2000 to the Regents of the University of Michigan (see Figure 4 for a screen shot of this EHC). Respondents in both conditions were asked about events that occurred over the life course.

EHC Questionnaire Design

Before administering the EHC, questionnaire development work on the life-course EHC involved two rounds of usability pretests with 10 subjects each. The pretests informed finalization of the EHC instrument, which included 9 time line domains (historical landmarks, residential, marriage and cohabitation, children, education, labor, parent/guardian, parental socioeconomic status, and health history), hierarchically organized to elicit temporal anchors during the early stages of the interview and then to collect more difficult-to-retrieve temporal information. Interviewers were given scripted questions to introduce each domain (see Appendix A) and were instructed on how to use top-down, sequential, and parallel cues. The scripts for the most part allowed respondents to begin any place within the life course, giving respondents the option of moving in forward, backward, or mixed chronological order. The historical landmarks domain did not elicit any information from the respondents. The landmarks domain included the dates of noteworthy events (e.g., assassination of JFK, 9/11) for interviewers to use as temporal anchors.

The residence domain was designed to collect information on any change in residence, and the addresses (street, city, state) of residences. Transitions in residence are another source of potentially useful anchors, primarily because such transitions may be related to changes in employment status. Interviewers were instructed to collect information that exhausted the life course timeline. This requirement maximizes the use of sequential retrieval, as respondents are encouraged to remember distinctive transition points exhaustively and what preceded and followed these transitions.

The marriage and cohabitation domain began with a question to determine the names or initials of spouses and partners. Respondents were then asked if each person was a spouse only, partner only or both. The marriage domain included four timelines where interviewers could record dates of living with spouses, periods of separation from spouses, spells of living with partners, and periods of separation from partners.

The children domain collected information about children fathered or given birth to and children legally adopted. Respondents were asked to give first names or initials for each child and asked to supply birth dates, adoption dates, and dates of death where appropriate. Respondents were also asked if each child ever lived apart from them for four months or more before they turned 18 years of age. A separate timeline was displayed for each child who lived away from the respondent.

The purpose of the education domain was to collect information on all of the respondent's academic experiences. The data entry matrix included a drop down menu to record everything from elementary school attendance through professional or graduate school.

The labor domain captured periods of employment over the life course, as well as annual employment status, full and part-time work status and spells of unemployment. In this domain, the respondent could be cued by being reminded of the timing of overlapping or contemporaneous jobs and of transition points that corresponded to residential changes.

The parent/guardian domain was designed to record those persons who were the respondent's live-in guardians from birth to age 18. Separate timelines were presented for nine possible guardian relationships: on own, mother and father, mother, father, stepmother, stepfather, grandmother, grandfather and other. If other was selected, interviewers were to record the person's relationship to the respondent.

The parental socioeconomics domain included two timelines. The first timeline was concerned with the employment and unemployment of the family's main breadwinner while the respondent was growing up. The second timeline recorded how well off the respondent thought their family was during their childhood.

The health history domain gathered information related to disabling health conditions, general health status, smoking and weight history.

Q-List Questionnaire Design

A parallel Q-list instrument was developed to measure the same key outcome variables as the EHC questionnaire. Pretests with 9 participants were conducted, with modifications to the instrument having been made on the basis of the pretests. The Q-list instrument included separate question sections to obtain comparable information from each domain outlined in the EHC design: places of residence, marriages and cohabitations, children, education, labor, parent/guardian history, parental socioeconomic status, and health history.

Because of the complicated nature of individual life histories, skip patterns were built into the questionnaire design to facilitate individual circumstances. For illustrative purposes, in what follows, examples are provided of possible questions in selected domains:

- Residence: What was your first address, what address did you live at when you were born? Have you always lived at <address>? (if no) In what month and year did you move from <address>? What was the address of the place that you moved to?
- Marriage: Now I'd like to ask you a few questions about marriage. Altogether, how many times have you been married? (if 2+) In what month and year did your first marriage begin? Did the marriage end in divorce, or were you widowed? In what month and year did the marriage end?
- Cohabitation: Have you ever lived with a partner as if married? (if yes) How many times have you lived with a partner as if married? (if 2+) In what month and year did you start living with a partner as if married for the first time? Did you ever marry this partner? (if no) In what month and year did you stop living with this partner?
- Employment: Are you currently working for pay? (if yes) What is the name of your employer, who do you work for? In what month and year did you start with <employer name> as your main job? Thinking back to when you first started working for <employer name>, were you working full or part time? (If full) Have you been continuously working full time since <date of start> until now? (if no) In what year did you change to working part time? We would like you to think of your most recent main job prior to your working at <employer name>. What was the name of your employer, who did you work for?
- Health Status: Now I'd like to ask about your general health. Thinking back to your early childhood, from birth until you reached the age of 7, would you say that your health was excellent, very good, good, fair, or poor? Since early childhood, has your health consistently stay at this level? (if no) In what year (at what age) did your health change?

Interviewer Training

Twenty-eight interviewers received a total of 21 hours of training over three days to conduct interviews. Interviewers were told that they were part of an experiment and were asked to refrain from discussing the study outside of their study group, however they were not informed about the other interviewing condition. Every effort was made to make the Q-list and EHC training sessions as complementary as possible.

EHC Interviewers were trained to use top-down, sequential, and parallel cues. They were also instructed to use narrowing probes. Upon exiting the interview, the EHC prompted interviewers to resolve inconsistencies in time line data and check for gaps in the time lines. An electronic message was also generated if the Interviewer failed to access any of the domains.

Q-list interviewers were trained to use a year/age conversion facility built into the standard computerized interview to help respondents determine a year if only an age was provided by the respondent, or an age if only a year was provided.

Interviewer Assessments

Following each interview, interviewers assessed their perceptions of respondents' cooperation, effort, ability to understand and answer questions, and ability to remember major life events, and detailed events. In addition, interviewers rated their own enjoyment, and ease of administration, of the interviews.

Results and Discussion

Results are preliminary, and they focus on two different aspects in their comparison of EHC and Q-list conditions. First, analyses of data quality examined the correspondence of the EHC and Q-list retrospective reports with panel reports that had been collected annually in the PSID from the same households as the respondents. The study design was such that the EHC and Q-list respondents provided the panel reports themselves concerning household conditions on at least half of the panel waves. As there is more retrospective information than panel information, these analyses are limited to those retrospective measures to which there exists adequate corresponding panel data. Second, analyses of operational measures were designed to test whether the EHC and Q-list conditions significantly differed with regard to interviewing time, and interviewer assessments on the quality of interviewer administrations.

Data Quality Analyses

Data quality analyses concentrate on the correspondence between experimental (EHC; Q-list) and panel reports for annual moves, annual amount of time employed, life course number of marriages, annual cohabitation, and annual health status.

Annual moves. Analyses to determine correspondence between annual retrospective reports and panel reports of moving are compromised because the panel data were not collected on a straightforward calendar-year basis. Instead, respondents were asked whether they had moved since the spring of the previous calendar year. Accordingly, the panel reports were scored by determining the probabilities of having moved in a given calendar year, on the basis of what was known about interviewing date (which varied by wave), and assuming that the start of Spring occurred on March 1. As for the retrospective reports, they were scored on the basis of whether there was no reported move in a calendar year, whether there was a single move within the same city (simple move), or whether there was a single move to another city, or more than two moves in a calendar year (regardless of city; complicated move).

First were examined the annual proportion of retrospective reports of not moving at all, or having simple or complicated moves; these proportions are reported in the top panel of Table 2. In comparison to the Q-list condition ($\underline{M} = .865$, $\underline{SD} = .085$), the EHC condition

led to reports of significantly fewer moves ($M = .848$, $SD = .081$; $F(1, 624) = 6.37$, $p = .01$), and to a significantly higher proportion of simple moves (Q-list $M = .042$, $SD = .045$; EHC $M = .063$, $SD = .042$; $F(1, 624) = 35.80$, $p < .0001$). The conditions did not significantly differ in the annual proportion of complicated moves. Next, to gain information on whether, for example, the elevated EHC reports of simple moves was due to overreporting, comparisons were made on the probability of an annual panel move for each annual retrospective report of having no move, a simple move, or a complicated move, and for each condition. If, for example, the panel probability that corresponds to a simple move in the EHC condition is significantly less than the probability observed in the Q-list condition, then respondents in the EHC condition, relative to those in the Q-list condition, would be overreporting simple moves. As indicated in the lower panel of Table 1, EHC and Q-list conditions did not significantly differ in the corresponding panel proportion of moving for no moves, simple moves, or complicated moves.

These results indicate that the EHC condition led to more accurate reporting of annual moves across the life span, but only when simple moves were made. For complicated moves, there were no differences between conditions. Thus, for more distinctive events, such as moves to a different city, or for more distinctive years, such as having moved more than once, the EHC and Q-list conditions did not differ. Yet, the EHC was able to encourage the recovery of less distinctive moves beyond what was available via Q-list interviewing (see also Means and Loftus, 1991; van der Vaart, in press; for similar results).

Annual amount worked. For each year, the panel data provide reports of the number of hours worked. For the retrospective reports, an annual score was calculated that consisted of a 6-point scale of amount worked based on reports of having worked the entire year, part of the year, or not at all that year, and whether one work entirely part-time, entirely full-time, or both part- and full-time while working. The data were equated between conditions, such that determinations of working the entire year or part of the year was based on reported months of working (and not based on the separate report of annual work status in the EHC condition). For each year, Pearson correlations between panel work hours and retrospective amount worked were calculated separately by condition, and the results are provided in Figure 5.

Figure 5 illustrates that for most years, the EHC condition provided higher Pearson correlations than did the Q-list condition. Using a goodness of fit test, the EHC condition did provide significantly more years ($n = 22$) of higher correlations in comparison to the Q-list condition ($n = 9$), $\chi^2(1) = 5.28$, $p = .02$. Interestingly, for recent years, the Q-list condition provided consistently higher correlations in comparison to the EHC, and thus the advantage of EHC interviewing appears restricted to more remote points in time. An explanation of these findings is one based on an order of retrieval effect. The Q-list always had respondents first think about their current jobs (if any) and then move backwards in time. Such an approach may have facilitated the remembering of more recent employment. Yet, overall, results are similar to those of Callegaro et al. (2004), who have found in a two-year reference period that reports of hospitalizations were more accurate in a Q-list condition for the more recent year, but that the EHC provided more

accurate reports for the more remote year, and thus may be more indicative of a general pattern of results with these conditions.

Number of marriages. The total number of marriages provided in the panel and retrospective reports were calculated for each respondent. These data were analyzed in three ways. First, signed differences were examined by subtracting the number of panel marriages from the number of experimental marriages; a Fisher's exact test found that EHC leads to significantly more underreporting in comparison to Q-list ($p = .012$). Next, the absolute value of the differences were examined between conditions and with a Fisher's exact test found that the EHC leads to significantly more absolute error ($p = .023$). Finally, compared kappa statistics were compared between conditions, and although both EHC ($\kappa = 0.84$) and Q-list ($\kappa = 0.92$) reports demonstrated excellent agreement with panel reports, the Q-list kappa is significantly stronger than in the EHC condition, $z = -2.22$, $p < .05$.

Results are unambiguous that the Q-list condition outperformed the EHC in reported number of marriages. An explanation of these results centers on the manner in which respondents were asked to report about their marriages. In the Q-list condition, respondents were directly asked how many times they were married; in the EHC condition, respondents were asked about the names or initials of all their partners, whether these relationships were marriages or not. Apparently, the number of times one has been married is, for most respondents, an easily retrieved piece of information that is reported accurately. Calendar-based interviewing that asks respondents directly how many times that they were married (see Freedman et al., 1988) may not show deficient reporting as observed in the EHC condition.

Annual cohabitation. In contrast to reports of marriage, results indicate that the EHC condition led to higher quality reports of cohabitation in comparison to the Q-list condition. For each year, panel and retrospective reports were compared on whether respondents indicated having a cohabiting partner or not, and agreement was measured via kappa indices. The results of these analyses are provided in Figure 6. Similar to results with annual amount worked, after assigning the 6 ties in kappa as 3 years to each condition, the EHC condition ($n = 18$) had significantly more years with higher kappas than did the Q-list ($n = 6$) condition as revealed by a goodness of fit test, $\chi^2(1) = 4.65$, $p = .03$. Contrary to the results with annual amount worked, the EHC condition is noted for having considerably higher kappas for the more recent years in comparison to the Q-list. Thus, there does not appear to be a general pattern in which the EHC provides consistently better data for the more remote years. Moreover, the manner of questioning used in the EHC, in which the names or initials of partners were elicited before the timing of the cohabitation spells were collected, appears to have benefited the quality of responses for cohabitation.

Annual health status. Both panel and retrospective reports collect annual reports of health status on a scale of 1 = excellent to 5 = poor. The panel data are constrained in that health status was not asked until 1984; thus analyses are based on reports for 1984 through 1997. To examine EHC and Q-list data quality, annual intraclass correlations

were first computed separately between conditions. For both conditions, intraclass correlations were consistently moderately strong as they ranged from .45 to .65. Although the EHC condition had more years ($n = 9$) in which there were higher intraclass correlations than the Q-list condition ($n=5$), a goodness of fit test did not yield a significant difference.

Next, a series of mixed-model analyses were performed to determine whether there were slope differences in the annual mean levels of health status between retrospective and panel (prospective) reports between EHC and Q-list conditions. The mean levels of health status across years are provided in Figure 7. Most noteworthy, the slopes for all observations are positive and significantly different from 0, indicating that respondents' were reporting worse health status as they aged both prospectively, and retrospectively. Interestingly, the slope of the Q-list retrospective data (slope = .015) is significantly shallower than the corresponding prospective data (slope = .024, $t(612) = 2.08$, $p = .04$). The slope of the EHC retrospective reports (slope = .021) did not significantly differ from the corresponding prospective data (slope = .022).

Again using a mixed-model approach, examined was whether predicted mean levels significantly differed between retrospective and panel (prospective) reports at the beginning year of the series (1984), at the midpoint (between 1990 and 1991), and at the ending year of the series (1997). For 1984, only the EHC condition had a lower predicted mean retrospectively in comparison to the corresponding panel predicted mean, $t(617) = 2.23$, $p = .03$. However, for both the midpoint, and for 1991, both EHC and Q-list conditions yielded predicted mean values that were significantly lower than those observed in their respective corresponding panel data (midpoint Q-list $t(618) = 2.42$, $p = .02$; 1997 Q-list $t(615) = 3.32$, $p = .001$; midpoint EHC $t(617) = 2.81$, $p < .01$; 1997 EHC $t(612) = 2.63$, $p < .01$).

In summary, respondents in both EHC and Q-list conditions were retrospectively reporting their health status as better in comparison to their prospective panel reports. Only the EHC, however, preserved the rate of change over time that was observed in the panel data. As for the Q-list, respondents reported less deterioration in health status over time retrospectively than observed in the corresponding panel reports.

Analyses of Operational Measures

Interviewing time. Belli, Shay, and Stafford (2001) observed no differences in interviewing time between paper and pencil EHC and Q-list questionnaires, that tested a two-year reference period. With the computer-assisted lifecourse EHC and Q-list instruments used in the present experiment, the EHC condition did lead to significantly longer interviews ($M = 57.6$ minutes, $SD = 28.7$) in comparison to the Q-list condition ($M = 51.5$, $SD = 21.9$), $t(615) = 2.96$, $p = .003$, although the 6-minute average difference between conditions is relatively modest in the context of a nearly hour-long interview.

Interviewer assessments.

Table 3 presents results from the interviewer assessments. Interviewers found respondents to be significantly more cooperative and motivated in the EHC condition in comparison to the Q-list. In addition, interviewers perceived respondents to be better able to understand and answer questions, and to remember both major and detailed past events, in the EHC condition. Finally, interviewers reported to have enjoyed the EHC interviews more, they found EHC interviews as being easier to administer and to record answers, and that the EHC computer application facilitated interviewing more than the Q-list application. These results largely replicate Belli et al.'s (2001) findings that interviewers preferred paper and pencil EHC interviews over Q-list ones.

Conclusion

The results of this paper continue to support the proposition that because of their ability to utilize a greater number of cues available in the structure of autobiographical memory, EHC interviewing yields higher quality retrospective reports than Q-list interviews. Although the EHC does not outperform the Q-list in each and every instance, the EHC does better more often, and consistent results have been found with shorter reference periods, and with different modes of interviewing (face-to-face, telephone, paper and pencil, computer-assisted). The implications of this body of research favor the notion that social and health scientists should seriously consider benefits that calendar-based interviews can provide to the quality of retrospective survey data to which they oft rely upon for their scientific inferences.

References

- Axinn, W. G., Barber, J. S., & Ghmire, D. J., (1997). The neighborhood history calendar: A data collection method designed for dynamic multilevel modeling. *Sociological Methodology*, 27, 355-392.
- Belli, R. F. (1998). The structure of autobiographical memory and the event history calendar: Potential improvements in the quality of retrospective reports in surveys. *Memory*, 6, 383-406.
- Belli, R. F., Lee, E. H., Stafford, F. P., & Chou, C-H. (in press). Calendar and question-list survey methods: Association between interviewer behaviors and data quality. Manuscript submitted for publication. *Journal of Official Statistics*.
- Belli, R. F., & Lepkowski, J. M. (1996). Behavior of survey actors and the accuracy of response. *Health Survey Research Methods: Conference Proceedings* (pp. 69-74). DHHS Publication No. (PHS) 96-1013.
- Belli, R.F., Shay, W. L., & Stafford, F. P. (2001). Event history calendars and question list surveys: A direct comparison of interviewing methods. *Public Opinion Quarterly*, 65, 45-74.
- Brown, Norman R., and Donald Schopflocher. 1998. "Event Clusters: An Organization of Personal Events in Autobiographical Memory." *Psychological Science* 9:470-75.
- Callegaro, M., Yu, M., Cheng, F-W., Hjermsstad, E., Liao, D., & Belli, R. F. (2004, May). Comparison of computerized event-history calendar and question-list

- interviewing methods: A two-year hospitalization history study. Paper presented at the annual AAPOR conference, Phoenix.
- Cannell, C. F., Fisher, G. & Bakker, T. (1965). Reporting of hospitalization in the Health Interview Survey. *Vital and Health Statistics* (PHS Publication No. 1000, Series 2, No. 6). Washington, DC: U.S. Government Printing Office.
- Caspi, A., Moffitt, T. E., Thornton, A., Freedman, D., Amell, J. W., Harrington, H., Smeijers, J., & Silva, P. A. (1996). The life history calendar: A research and clinical assessment method for collecting retrospective event-history data. *International Journal of Methods in Psychiatric Research*, 6, 101-114.
- Dugoni, B., Lee, L., & Tourangeau, R. (1996). *Report on the NLS/Y round 16 recall experiment*. Unpublished manuscript: NORC.
- Duncan, G. J., & Hill, D. H. (1985). An investigation of the extent and consequences of measurement error in labor economic survey data. *Journal of Labor Economics*, 3, 508-522.
- Duncan, G. J., & Mathiowetz, N. A. (1985). *A validation study of economic survey data*. Ann Arbor, MI: University of Michigan.
- Furstenberg, F. F. Jr., Brooks-Gunn, J., & Morgan, S. P. (1987). *Adolescent mothers in later life*. New York: Cambridge University Press.
- Freedman, Deborah, Arland Thornton, Donald Camburn, Duane Alwin, and Linda Young-DeMarco. 1988. "The Life History Calendar: A Technique for Collecting Retrospective Data." In *Sociological Methodology* (Vol. 18), ed. C.C. Clogg, pp. 37-68. San Francisco; Jossey-Bass.
- Kessler, R. C., & Wethington, E. (1991). The reliability of life event reports in a community survey. *Psychological Medicine*, 21, 723-738.
- Lyketsos, C. G., Nestadt, G., Cwi, J., Heithoff, K., & Eaton, W. W. (1994). The life chart interview: A standardized method to describe the course of psychopathology. *International Journal of Methods in Psychiatric Research*, 4, 143-155.
- Mathiowetz, N. (1986). The problem of omissions and telescoping error: New evidence from a study of unemployment. *Proceedings of the Section on Survey Research Methods*, American Statistical Association.
- Mathiowetz, N. A., & Duncan, G. J. (1988). Out of work, out of mind: Response errors in retrospective reports of unemployment. *Journal of Business and Economic Statistics*, 6, 221-229.
- Means, B., & Loftus, E. F. (1991). When personal history repeats itself: decomposing memories for recurring events. *Applied Cognitive Psychology*, 5, 297-318.
- Schank, R. C., & Abelson, R. P. (1995). Knowledge and memory: The real story. In R. S. Wyer (Ed.), *Advances in social cognition*, Vol. VII (pp. 1-85). Hillsdale, NJ: Erlbaum.
- Van der Vaart, W. (in press). The time-line as a device to enhance recall in standardized research interviews: a split ballot study. *Journal of Official Statistics*.
- Yoshihama, M., Gillespie, B., Hammock, A. C., Belli, R. F., & Tolman, R. M. (2004). Does the Life History Calendar Method Facilitate the Recall of Domestic Violence Victimization? Manuscript submitted for publication.

Figure 1. Percentage reported moving during 1996: Study 1.

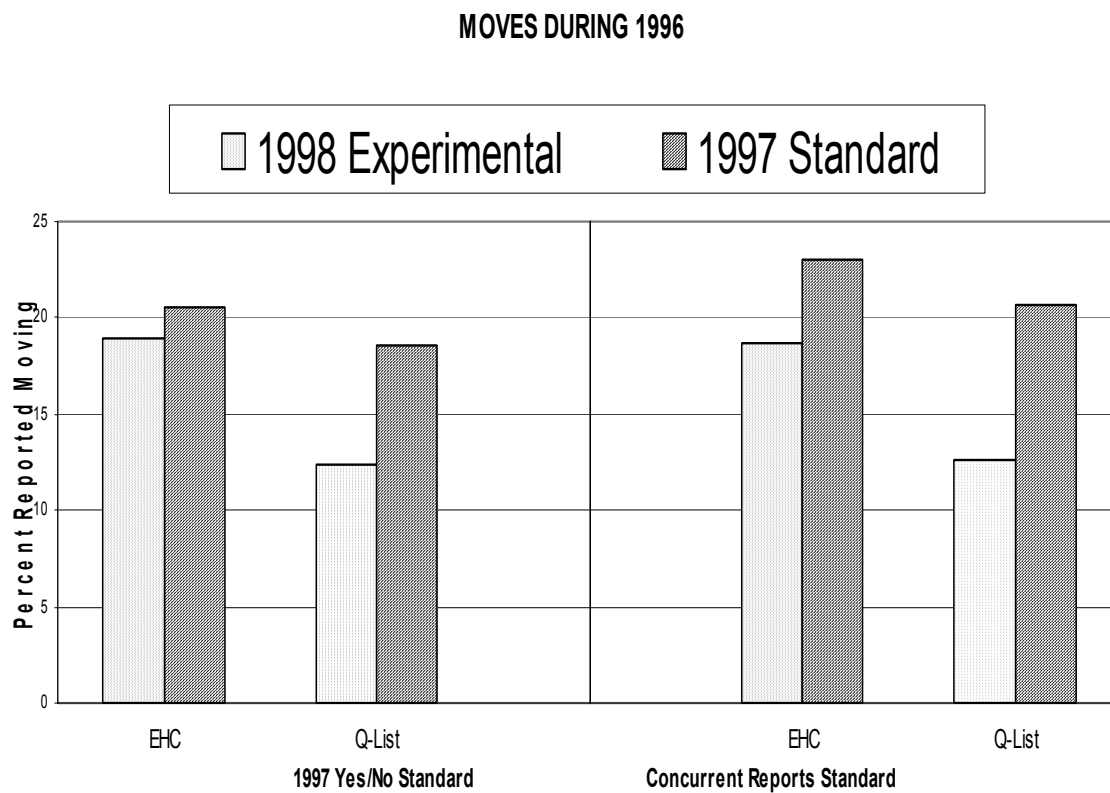


Figure 2. Number of jobs held during 1996: Study 1.

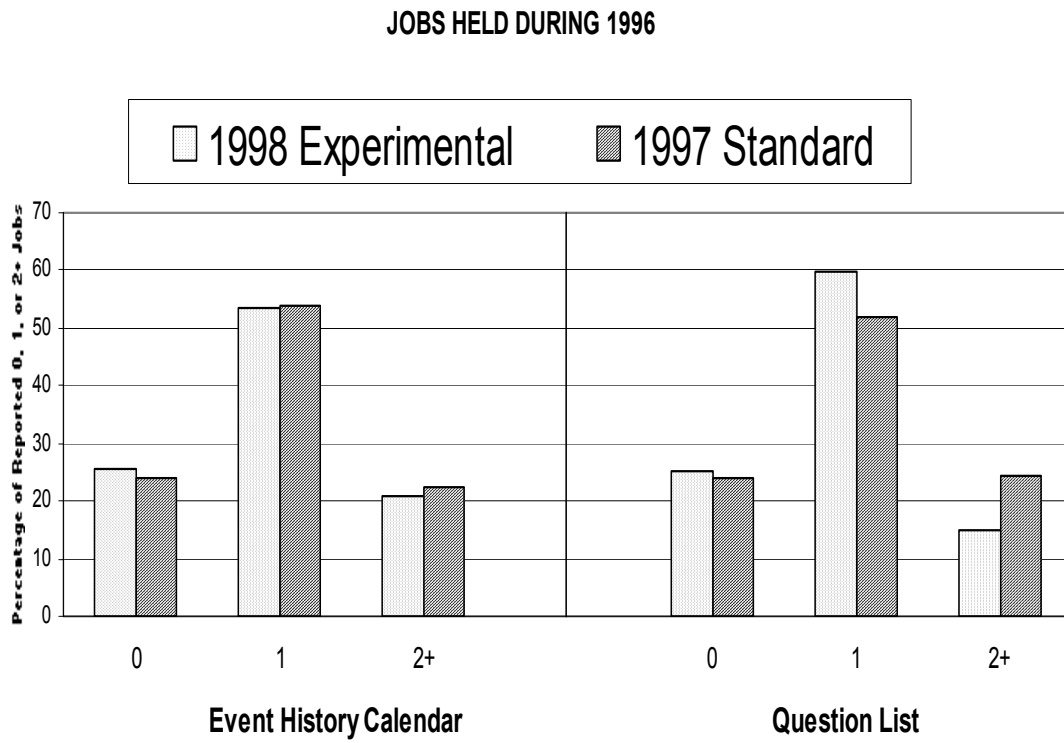


Figure 3. Pearson correlations between experimental (EHC; Q-list) and standard of comparison reports for different outcome measures: Study 1.

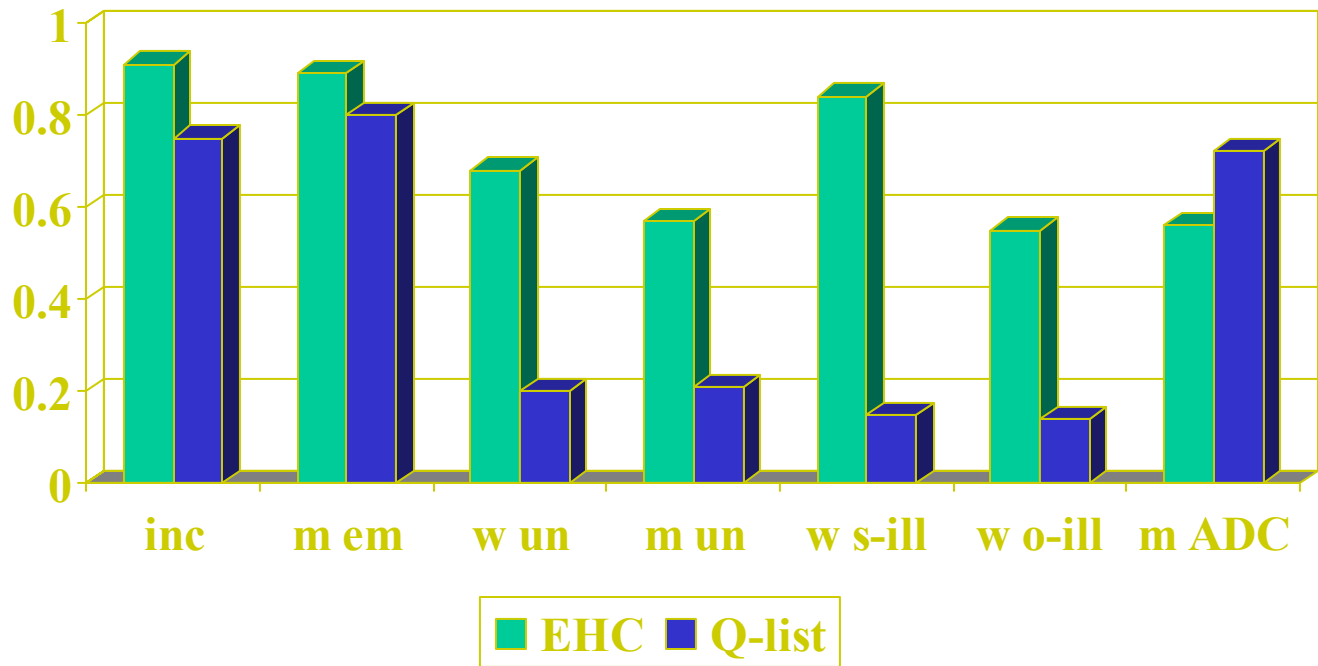


Table 1. Prevalence of Reported IPA: Study 2

Group	Percent Reported IPA				Percent estimated to experience any type by age 50
	Any Type	physical	sexual	threat	
Q-list	63.5	59.3	27.0	38.7	73.2
EHC	92.5	87.5	52.5	70.0	96.1

Table 2. Percent annual retrospective reported moves, and probability of corresponding prospective move, for EHC and Q-list conditions: Study 3.

	No move	Simple move	Complicated
% annual retrospective reported moves			
EHC	84.8	6.3	8.9
Q-list	86.5	4.2	9.3
p	= .01	<.0001	n.s.
Probability of corresponding prospective move			
EHC	.100	.494	.607
Q-list	.097	.540	.580
p	n.s.	n.s.	n.s.

Figure 4. Screen capture of the lifecourse EHC for the labor history domain. The EHC is copyright (c) 2000 to the Regents of the University of Michigan.

The screenshot displays the Event History Calendar (EHC) for the labor history domain. The interface is divided into several sections:

- Event History Calendar (EHC) - [Employer]**: The top section shows a timeline from 1960 to 1990. The timeline is divided into four segments: 1960-1969, 1970-1979, 1980-1989, and 1990-1999. Each segment has a color-coded header (blue, green, yellow, and red respectively) and a grid of years (0-9) for data entry.
- Residence**: A section below the calendar showing a timeline of residence data. A tooltip indicates the address: "1980, 1982, 333 Nice Place, Davenport, IA".
- Children and Marriage**: A section below the residence data showing a timeline of children and marriage data.
- Main Employer**: A section below the children and marriage data showing a timeline of main employer data.
- Labor Data Entry Window**: A section below the main employer data showing a timeline of labor data entry. It includes fields for:
 - Employer**: A grid of years (0-9) for data entry.
 - Annual Employment Status**: A grid of years (0-9) for data entry.
 - Full/Part-time work**: A grid of years (0-9) for data entry.
 - Unemployment**: A grid of years (0-9) for data entry.
- Navigation and Footer**: The bottom of the window features a navigation bar with tabs for Residence, Marriage, Children, Education, Labor, Parent/Guardian, Parental SES, and Health. Below the navigation bar is a button labeled "Employer Questions". The Windows taskbar at the bottom shows the Start button, the EHC application, and the Microsoft Word - HE... document. The system clock indicates 8:50 AM.

Figure 5. Correlations between retrospective amount worked and prospective work hours: EHC and Q-list conditions: Study 3.

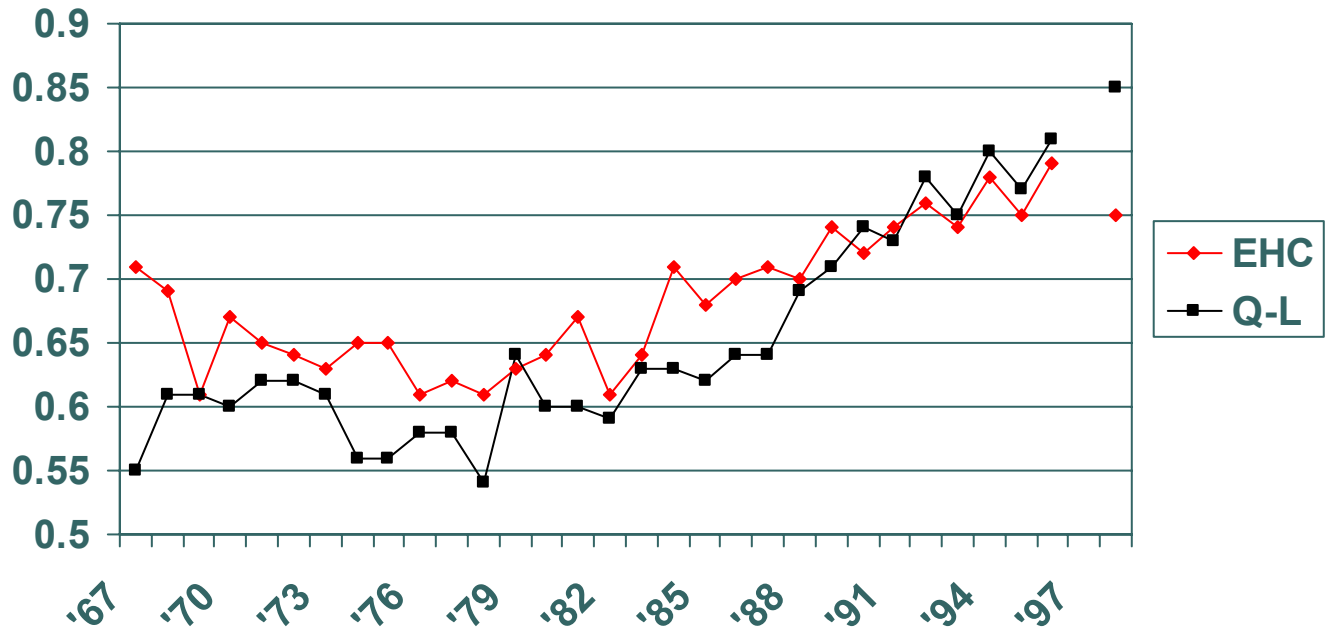


Figure 6. Kappas between retrospective and prospective cohabitation partners: EHC and Q-list conditions: Study 3.

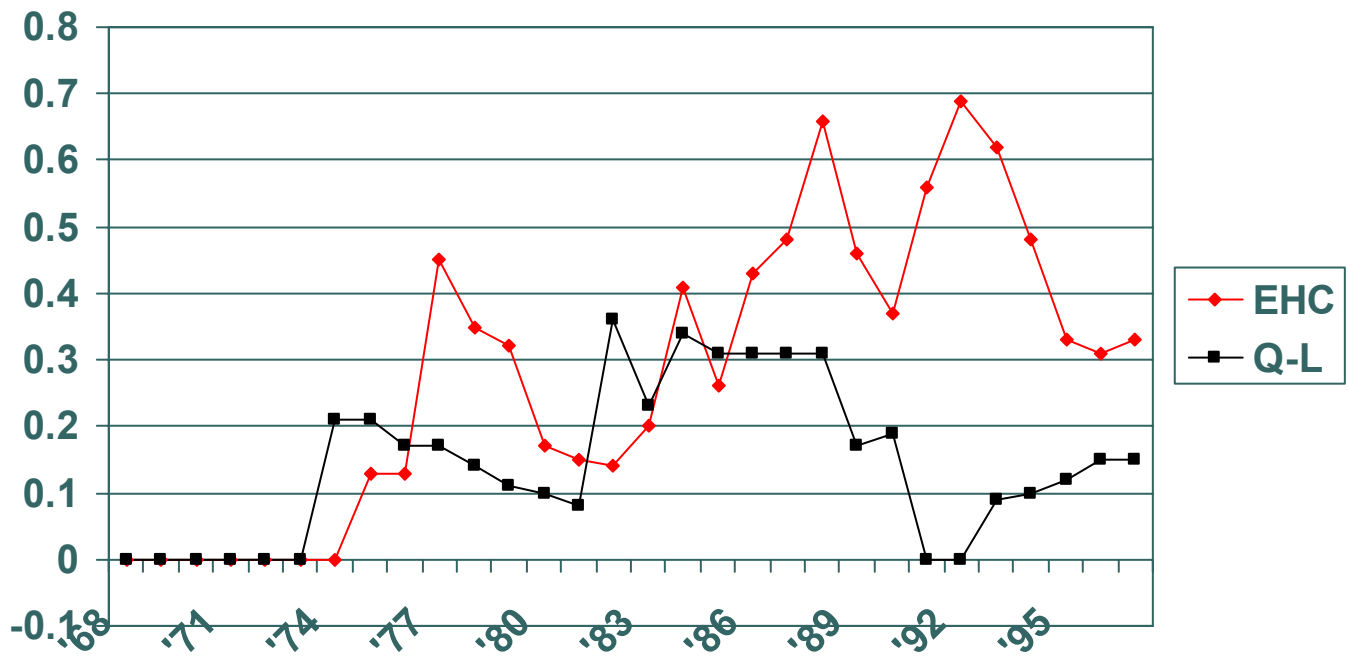


Figure 7. Mean prospective and retrospective health status (1 = excellent; 5 = poor):
EHC and Q-list conditions: Study 3.

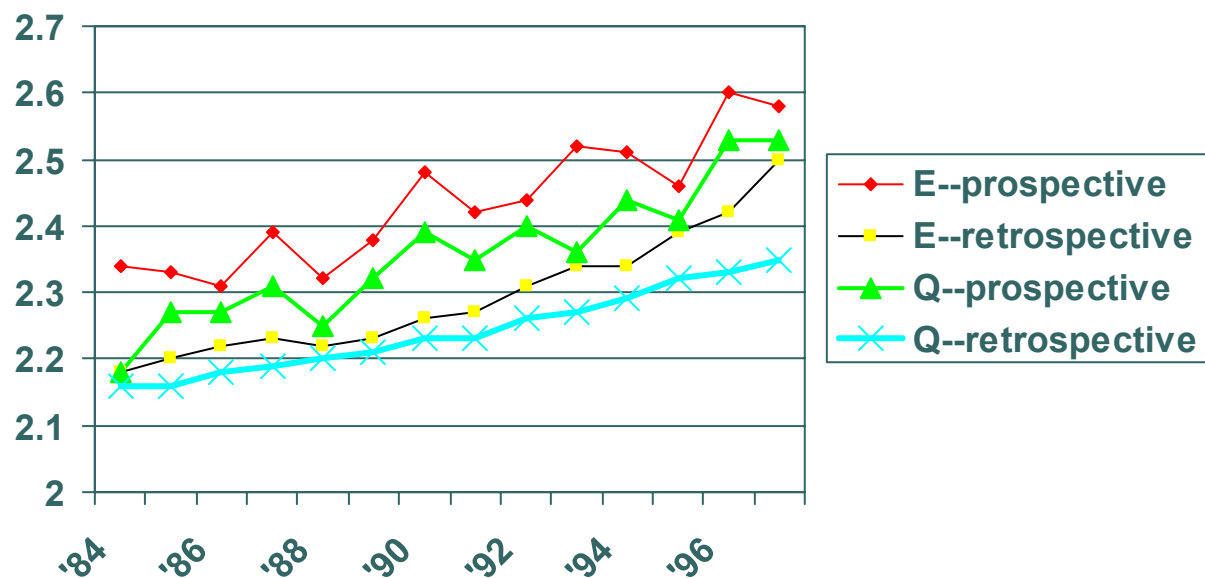


Table 3. T-tests to determine mean difference in interviewer observations between EHC and Qlist interviewing conditions: Study 3.

Interviewer Observation	EHC Mean (sd)	Qlist Mean (sd)	t-test	p
How was the respondent's cooperation during the interview? (1=excellent to 4=poor)	1.17 (.45)	1.40 (.62)	5.10	<.0001
How much effort did the respondent put into answering the questions? (1=a lot to 4=none)	1.21 (.43)	1.37 (.58)	3.81	.0002
How would you rate how much you liked or enjoyed the interview with this interview? (0=very much disliked to 6=very much enjoyed)	6.13 (1.25)	5.75 (1.29)	3.74	.0002
How would you rate how difficult or easy it was to administer the questions during the interview? (0=very difficult to 6=very easy)	5.94 (1.31)	5.07 (1.66)	7.15	<.0001
How would you rate how difficult or easy it was to record the responses during the interview? (0=very difficult to 6=very easy)	5.71 (1.49)	5.10 (1.51)	5.06	<.0001
To what extend did your working with the computer program hamper or facilitate interviewing this respondent? (0=hampered a great deal to 6=facilitated a great deal)	5.43 (1.47)	5.03 (1.31)	3.61	.0003
How would you rate the respondent's ability to understand and answer the questions that were asked? (0=not at all to 6 very able)	6.12 (1.22)	5.60 (1.30)	5.10	<.0001
How would you rate the respondent's ability to remember major life events? (0=not at all to 6=very able)	5.79 (1.38)	5.32 (1.47)	4.08	<.0001
How would you rate the respondent's ability to remember more detailed events? (0=not at all to 6=very able)	5.37 (1.52)	4.84 (1.69)	4.05	<.0001

Appendix A. Health and Economic Measures: EHC Domain Scripts: Study 3

Residence

Let's start with places that you have lived. Can you tell me the addresses of each of the places that you have lived over your entire lifetime and when you lived there? You can start since when you were growing up or you may want to think of where you are living now and work backwards in time.

Of course I'd be interested in where you were living when you were born, and any moves that you made when you were very young. You may have been too little to remember some moves directly, but you may know about them from what other family members have told you.

Marriage

(For Names that are both a spouse and partner): Now I'm interested in years in which you were living with NAME while legally married, and those years in which you were living with NAME while not married. I am also interested in any times in which you were living with NAME for 4 months or more.

(For Name that is only a spouse): Now I am interested in those years in which you were living with NAME while legally married. I am also interested in any times in which you were not living with NAME for 4 months or more for any reason.

(If stopped living with spouse before the end of the reference period): Did you stop living with NAME because of separation, divorce, or because you were widowed?

(For Names that are only a partner): Now I am interested in those years in which you were living with NAME as if married. I am also interested in any times in which you were not living with NAME for 4 months or more for any reason.

Children

So that I can keep track of your children as we talk about them, can you please tell me the first name or provide initials of each child you (fathered/gave birth to) or formally adopted.

While NAME was growing up, from (birth/time of adoption) until (turning 18 years of age/the time of death) was there ever a period of 4 months or more when (NAME) was living apart from you?

Education

I am now interested in the formal education that you have had over your entire lifetime. Please tell me about those periods in which you were attending elementary school, middle or junior high school, and high school. If you did not graduate from high school, but took classes to earn a GED, I would like to know about this as well.

In addition, I would like to know if you attended college, and professional and graduate school. For these schools, I would like to know whether you were attending part- or full-

time, based on the number of credits that were taking. (Please do not include trade or vocational schools such as beauty school, barber college, and so on).

Labor

Main Employers: Now I would like to talk with you about your work for pay. Have you ever worked for pay at the same job for 3 months or more? (If yes): Could you please tell me about your jobs for pay since you were 14 years of age, or when you began working, including any self-employment you may have had?

Annual Employment Status: So, when you were working for [Employer Name] for each year were you working the entire year or just part of the year? If you were working just part of the year for [Employer Name], did you have any other jobs that year so that you did have a job for the entire year?

Full-Time/Part-Time: When you were working for [Employer Name] from [start date] to [end date], was that full-time, part-time or some of both? (If part-time): Did you work at any additional jobs so that all together it was full-time?

- Remember that in any given year, you could have been working both full- and part-time.
- Full-time=30 hours a week or more

Unemployment: Finally, I am also interested in knowing about any periods of time of one week or longer since you were 14 years of age in which you did not have a job for pay, and you were looking for work at that time.

Parent/Guardian

During their lives, people often have different kinds of family arrangements, wherein children or teens are cared for by parents, stepparents, grandparents, or other adults. The next several questions are about the time when you were growing up, so please think back to when you were a young child and a teenager. How old were you when you moved out on your own? Who was responsible for taking care of you (before then/until you turned 18 years of age)?

For each person who served as primary parent/guardian: What was the highest grade of school or year of college that (he/she/they) completed?

Did (he/she/they) ever work for pay? (If yes): What was the most important job that (he/she/they) held?

Parental Socioeconomic Status

Main Breadwinner Employment/Unemployment: During the period in which you were growing up, from when you were born until you (moved out on your own/turned 18 years of age), we would like to know during which years the main breadwinner of your family was working and not working.

(If the main breadwinner was working for most of those years, ask: while you were growing up, did the main breadwinner ever lose his/her job and not find one right away?)

(For periods during your very young childhood, you may know about your family's main breadwinner's work from what other family members have told you.)

Family financial well-being: During each year of your childhood, from when you were born to 18 years of age, how well off was your family? Was it very well off, above average, average, below average, or very poor? It may help to focus on those years in which there was a change in how well-off your family had been. (For periods during your very young childhood, you may know about how well off your family was from what other family members have told you.)

Health

Disabling Health Conditions: We would like to know about any instances in which, because of injury, illness or disability you missed attending school or work for one month or more. Please tell me when any of these periods happened, how long they occurred and the type of injury or illness that it was. For periods during your very young childhood you may know about periods of injury or illness from what other family members have told you. We are also interested in any periods of time during your lifetime in which, because of injury, illness, or disability, you were confined to a hospital or to a bed at home for one month or more.

Health Status: We would also like to know how your general health has been over your entire lifetime. For each year of your life, would you say that your health had been excellent, very good, good, fair or poor? It may help to focus on those years in which there was a change in your general health from one state to another.

Smoking: Did you ever smoke cigarettes? (If yes): How old were you when you first started? We are interested in those times in your life when you did not smoke at all, when you smoked between 1-10 cigarettes per day, between 11-20 cigarettes a day, and 21 or more per day.

Weight: Finally, I'd like to ask you about your weight. Please think of five different categories that may have applied to you at different times during your life, you were very overweight, slightly overweight, just right, slightly underweight, or very underweight. So, during any period of your life were you very overweight? Slightly overweight? Just right? Slightly underweight? Very underweight?